

Syntactic, or Lexical, Zero in Natural Language

Author(s): Igor A. Mel'čuk

Proceedings of the Fifth Annual Meeting of the Berkeley Linguistics Society (1979), pp. 224-260

Please see "How to cite" in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Syntactic, or Lexical, Zero in Natural Language*

Igor A. Mel'čuk
Université de Montréal

This article is an attempt to answer some of the questions posed by A.A. Xolodovič [1] concerning the elaboration of a consistent theory of grammatical voice and the description of voice systems in natural languages. The following are three of his questions.

Question 1. What is the relationship of voice to such constructions as the Russian indefinite-personal agent construction of (1)

- (1) Ivana priglasili k trem časam
I. invited to three o'clock
(acc) (3pl)
'Ivan was invited for three o'clock'
(lit. '(they) invited Ivan...')

or the Russian impersonal construction of (2)?

- (2) Ivana oprokinulo
I.(acc) knocked over(neut sg)
'Ivan got knocked over' (lit. '(it) knocked
Ivan over')

Question 2. Are constructions (1) and (2), or examples such as (3), in Ukrainian,

- (3) Bulo organizovano ekspedyciju
was organized expedition
(neut sg) (neut sg) (acc)
'an expedition was organized'

grammatically subjectless? The same question may be asked of the following Spanish construction:

- (4) Aquí se vende periódicos
here sell newspapers
'newspapers are sold here' (lit. 'here
(it) is-sold newspapers')

Question 3. For Russian sentences such as (5)

- (5) Ivan byl priglašen k trem časam
Ivan was invited to three o'clock
(nom) (p part, masc sg)
'Ivan was invited for three o'clock'

why do we speak of an agentless passive, rather than positing an unspecified personal agent? Clearly (5), the true passive, is perfectly synonymous with (1), and (1) is a classic example of the unspecified personal agent construction (in Russian traditional grammar, indefinite personal).

To respond to these questions requires, first of all, clarification of the notion 'syntactic, or lexical, zero'. This in turn presupposes a theoretical account of linguistic-1 zero in general. [2] More generally, it presupposes an account of sentential incompleteness, which includes ellipsis (deletion of some of the lexical nodes), missing arguments, unfinished utterances, etc. Since these problems could not be fully investigated in a short paper, I will restrict myself here to the problem of zero in syntax. In the absence of a unified theory of syntactic incompleteness, my conclusions are of a preliminary and programmatic character.

To avoid overburdening the exposition I have refrained from reviewing the question and have provided only minimal references. For a detailed presentation of the problem of zeroes, and copious references, see Meier 1961; my main source of inspiration is the excellent analysis of Haas 1957. On zero affixation in word derivation see Lopatin 1966, Kastovsky 1969; remarks pertinent to zero expression of categories are found in Aschmann and Wonderly 1952. An insightful analysis of zeroes in syntax is provided by Wierzbicka 1966.

I will be working within the linguistic-2 framework of meaning-text theory, the basic assumptions of which I will presume the reader knows (see Žolkovskij and Mel'čuk 1967, 1969; Mel'čuk 1967, 1972, 1974a, 1976:26-62). I adopt the following plan of exposition.

In section 1, syntactic, or lexical, zeroes such as the zero wordform or (better) zero lex (see fn. 10) and zero lexeme are introduced as generalizations of such morphological zeroes as zero morphs and zero morphemes. The notions of zero morph and zero morpheme are sufficiently clear and intuitively accessible to be taken as basic, at least in certain instances, e.g., the zero genitive plural ending of Russ. spin-0 'of (human) backs', plit-0 'of slabs'; the zero masculine singular past ending of vstal-0 '(he) arose', mog-0 '(he) was able'; the zero masculine singular ending of zdorov-0 '(he is) healthy', vzjat-0 '(he is) taken'. All these zeroes contrast with non-zero endings elsewhere in the respective paradigms.

In section 2, the concepts of zero lex and zero lexeme are tested on certain types of Russian sentences.

In section 3, I discuss a use of the term zero verb (the notion is logically a particular instance of the notion 'syntactic, or lexical, zero') current in linguistic-2, par-

ticularly Russian, literature; the goal is an improved organization of terminology for zeroes. Some of the uses of this term correspond not to the notions I propose but to phenomena sharply distinct from zero - in particular, ellipsis. As early as half a century ago Bally (1922) insisted that ellipsis must be strictly distinguished from syntactic zero (see also Vardul' 1969).

In section 4, answers to Xolodovic's questions are proposed, and several connections are noted between the theory of grammatical voice and the approach to zero lexes and lexemes outlined here.

1. A generally recognized example of a morphological zero is the zero ending in such Russian wordforms as /ruk/ 'of hands' (cf. nom. sg. ruka 'hand'). The ending is a zero morph: ruk-0 - gen. pl.

The term morph is to be understood as follows. A morph X is an elementary linguistic-1 sign, i.e., an ordered triple of the form: [3]

(6) X = < /X/; 'X'; x >, where:

/X/ is the signifiant of the morph, i.e., a string of phonemes (possibly with a complex of prosodemes), or a string of graphemes if a written text is considered; 'X' is the signifié of the morph, i.e., the symbol or complex of symbols which represents its meaning; and x is the syntactics of the morph, i.e., the set of all necessary data about the combinatorial possibilities of the morph X with other morphs. Syntactics may include cooccurrence restrictions based on part of speech, grammatical gender, phonological and/or morphological environment; selectional restrictions of all kinds; etc.

For example, the two non-zero genitive plural endings of Russian:

(7) -ěj = < /ěj/; 'pl., gen.'; \acute{Z} = nominal ending, second declension, non-neuter gender, not after palatalized or hushing consonants,...>;
-ov = < /of/; 'pl., gen.'; \acute{Z} = nominal ending, second declension, non-neuter gender, not after palatalized or hushing consonants,...>

Examples: /kon' + ěj/ konej 'of horses', /krovat' + ěj/ krovatej 'of beds', /mor' + ěj/ morej 'of seas', /nož + ěj/ nožej 'of knives'; vs. /stol + of/ stolov 'of tables', /tigr + of/ tigrov 'of tigers', /treugol'nik + of/ treugol'nikov 'of triangles', etc.

- (8) Eng. -d = < /d/; 'past'; ζ = verbal ending, in weak verbs, after vowel or voiced consonant other than /d/,...>;
-ed = < /d/; 'past'; ζ = verbal ending, in weak verbs, after /d/ or /t/,...>
-t = < /t/; 'past'; ζ = verbal ending, in weak verbs, after voiceless consonants other than /t/, or in specified strong verbs,...>

Examples: echoed, cared, dubbed, raged, struggled,...; faded, patted,...; missed, stopped, kicked, slept, meant,...

A zero morph is a morph whose signifiant is the empty string (of phonemes or graphemes):

- (9) Russ. \emptyset - pl, gen = < Λ ; 'pl., gen.'; ζ = nominal ending, first or second declension; not after palatalized or hushing consonant (if the noun is masculine, further specification is required)...>

Like any non-zero morph, a zero morph has both a signifié and syntactics. The signifié in the case of (9) coincides with that of the non-zero morphs of (7); the syntactics provides for the fact that the zero morph of (9) is restricted to certain types of Russian nominal stems.

Morphs having identical signifiés and sufficiently similar distributions (stated with phonological and/or morphological rules) are united into one morpheme. Thus the English morphs of (8) are allomorphs of the morpheme {PAST}; the Russian morphs -ej, -ov, and \emptyset - pl, gen are allomorphs of the morpheme {PL. GEN}, \emptyset - pl, gen being the zero allomorph of this morpheme. [4]

It is well known that there are morphemes with only one allomorph. For example, in Russian the morpheme {PL. LOC} for adjectives contains only one allomorph, -ix. The sole allomorph of a morpheme can also be zero: a morpheme that has only a zero allomorph is naturally called a zero morpheme, e.g., the nominal and adjectival singular morpheme in Spanish, {SG}, contains just one allomorph, zero. [5]

It follows that the expression 'morphological zero' can be understood in one of two precise senses: either (i) it is a zero morph, i.e., a morph whose signifiant is an empty string; or (ii) it is a zero morpheme, i.e., a morpheme having zero as its sole allomorph. We may make the following generalization:

- (10) A linguistic zero either is a zero sign, i.e., a sign whose signifiant is an empty string; or it is the corresponding zero '-eme', i.e., the set of equivalent signs which contains only one zero 'allo-' (one zero sign).

However, the morphological considerations used so far and the syntactic considerations of interest here are two very different things. In order to relate them we must resort to a more abstract level of analysis.

In morphology, the set of wordforms is considered given. A wordform is generally a complex, i.e., non-elementary, linguistic-1 sign: the signifiant of a wordform is a string of phonemes or graphemes (plus, probably, prosodies), and its signifié is a formal expression. [6] The deep morphological representation of a wordform used in the meaning-text model is such a formal expression; it consists of the name of the lexeme and a notation of morphological categories. (The syntactics of the wordform is irrelevant to the present discussion and can be safely omitted.) The goal of morphology is to describe the set of all wordforms of a language in a more concise manner than by simply listing them. This goal can be achieved to the extent that wordforms are complex signs, i.e., insofar as their signifiants have recurring components corresponding to recurring components in their signifiés. Here is one generally accepted approach to the construction of such a morphological description.

Consider the following set of signifiants of Russian wordforms:

- (11) /ruk/ 'of hands', /st'en/ 'of walls', /l'in'ij/ 'of lines', ...
 /nocej/ 'of nights', /put'ej/ 'of ways', /nozej/ 'of knives', ...
 /stvolof/ 'of (gun) barrels', /ostrovof/ 'of islands', /bojcof/ 'of soldiers', ...

and the corresponding set of deep morphological representations depicting their signifiés:

- (12) RUKA-pl, gen STENA-pl, gen LINIJA-pl, gen
 NOČ'-pl, gen PUT'-pl, gen NOŽ'-pl, gen
 STVOL-pl, gen OSTROV-pl, gen BOJEC-pl, gen

We must establish correspondences between sets (11) and (12) in the most economical way possible, without simply recopying all the pairs. To accomplish this the following morphs are introduced:

- (13) ruk = < /ruk/; 'RUKA'; Σ = stem, noun, fem, I decl,...>;
sten = < /st'en/; 'STENA'; Σ = stem, noun, fem, I decl,...>;...

ostrov = < /ostrov/; 'OSTROV'; = stem, noun, masc,
II decl,...>;
bojec = < /bojec/; 'BOJEC'; = stem, noun, masc, II
decl,...>

-Ø-pl, gen = < ; 'pl, gen'; = nominal ending, ei-
ther I decl or...> (see (9) above);
-ej = < /ej/; 'pl, gen';...> (see (7) above);
-ov = < /of/; 'pl, gen';...> (see (7) above)

The morphs of the language described must be given in a list. Obviously this list will contain far fewer items than the full list of wordforms. (For instance, if we take into account regular word formation rules of Russian, the list of Russian wordforms is at least 100 times the size of the list of morphs.) The correspondence of (11) to (12) can be described in a trivial manner: in the transition from (11) to (12), the signifiant of the wordform is broken into signifiants of the appropriate morphs in accordance with their syntactics; then the signifies of these morphs are combined to produce the deep morphological representation of the wordform. Conversely, in the transition from (12) to (11) the underlying deep morphological representation of the wordform is broken into the signifies of the morphs in accordance with their syntactics, then the signifiants of these morphs are combined to produce the signifiant of the wordform. (This description of the correspondence between signifiants of word forms and their deep morphological representation is very tentative. Accentuation and morphophonological alternations have not even been discussed.) By following this approach we arrive at an intermediate representation of wordforms - a representation in terms of morphs. Thus from a morphological point of view we need at least three levels for representation of wordforms:

- an upper (n + 1th) level - the signifiant of a wordform, a string of either phonemes or graphemes;
- an intermediate (nth) level -- a string of morphs;
- a lower (n - 1th) level -- the deep morphological representation of the wordform

For example:

n+1 level	/noč	éj/	/stvol	óf/	/ruk/	
	⇕	⇕	⇕	⇕	⇕	⇕
nth level	noč + ej		stvol + ov		ruk + ø-pl, gen	
	⇕	⇕	⇕	⇕	⇕	⇕
n-1 level	NOC-pl,gen		STVOL-pl,gen		RUKA-pl,gen	

An upper level is (closer to) text, a lower level (closer to) meaning. The double arrow symbolizes correspondence in both directions.

This schema makes explicit the distinctive formal property of zero morphs: a zero morph is the symbol in an intermediate representation of the nth level such that:

in the transition $\underline{n} \implies (n-1)$ it corresponds to some (usually non-empty) symbol or complex of symbols, while

in the transition $\underline{n} \implies (n+1)$ it corresponds to an empty string; and on its own nth level it combines with other symbols of the same level according to fully specified rules of the language which are captured in its syntactics.

A similar formal property must be inherent in any zero sign in language: a zero sign is a symbol of the nth level of representation which always corresponds, on the n+1, more nearly surface, level to an empty string and which is justified either from below (it corresponds to non-zero symbols on the n-1 level) or on its own level (it is governed by well-specified rules of cooccurrence with other symbols of level n.) [7]

In meaning-text theory and in many other systems another level of morphological representation is postulated: here, n'. The n' level is the '-emic' level lying between n and n-1. At the n' level a wordform is represented as a set of morphemes: {NOC'} + {PL. GEN}, {STVOL} + {PL. GEN}, {RUKA} + {PL. GEN}. (Recall that a morpheme is a set of morphs. Thus {NOC'} = [noč ~ nač], {STVOL} = [stvol ~ stvol' ~ stval ~ stval'], and {RUKA} = [ruk ~ ruk' ~ ruč]. The variant morphs represent Russian morphophonemics.)

Now we can construct syntactic analogs of zero morphs and zero morphemes. Note, however, that the analogy between syntax and morphology cannot be complete: the differences between the two are so fundamental that the parallels I will draw are little more than heuristic considerations.

Before discussing the zeroes of syntactic level \underline{n} we must determine the $\underline{n+1}$ and $\underline{n-1}$ levels of syntactic representation. I emphasize that from the viewpoint of the problem of zeroes these levels must be taken as axiomatic - the levels $\underline{n+1}$ and $\underline{n-1}$ must be determined independently of and prior to an inquiry into the question of zero on the \underline{n} level. In syntax the set of sentences is considered as given (thus a sentence is to syntax what a wordform is to morphology). A sentence is a complex of linguistic signs, one which has a signifiant and a signifie but lacks syntactics. [8] The signifiant of a sentence is a phonemic string plus relevant prosody (intonation contour, stresses, pauses); it constitutes the upper, or $\underline{n+1}$, level. [9] The signifie of a sentence - its semantic representation - is a graph introduced to represent the meaning of the sentence; it is the lower, $\underline{n-1}$, level. The syntactic analogs to morphs, the building blocks of the wordform, are wordforms, the building blocks of the sentence. Henceforth, I will use the term lex, which is not fully synonymous to wordform, but is preferable here in that it makes explicit the proportion morph : morpheme = lex : lexeme. [10]

Therefore an intermediate representation on the level \underline{n} in syntax is a sequence of lexes, each of them represented by its respective deep morphological representation. Thus for (14)

- (14) Fedor snova polez na mačtu
 F. again began to climb on mast
 'Fedor began to climb the mast again'

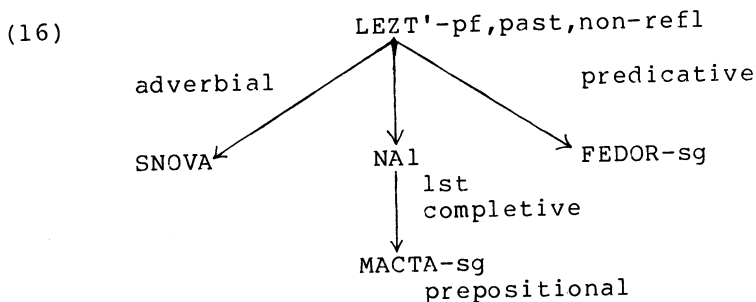
the representation of level \underline{n} in syntax will be

- (15) FEDOR-sg,nom SNOVA LEZT'-pf,past,non-refl,sg,m NAL
 MAČTA-sg,acc [11]

On the level \underline{n} , i.e., in terms of syntactic '-emes', the sentence is represented by surface syntactic structure. The surface syntactic structure of a sentence is a dependency tree whose nodes are not ordered from left to right and which stand in a one-to-one correspondence with all the lexes (including zero lexes) of the sentence. Each node is labelled with the name of the lexeme to which its lex belongs. The name of the lexeme is provided with indices of all the meaning-bearing morphological characteristics of the lex: number in nouns; tense, aspect, and mood in verbs. (Syntactically conditioned, i.e., non-semantic, morphological characteristics such as case in nouns and person, number, or gender in verbs, are omitted from surface syntactic structure. Each branch of the tree

is labelled with the name of the corresponding surface syntactic relation. The surface syntactic relations make explicit the interword connections that are ordinarily expressed in the language by such devices as word order and 'syntax-oriented' morphology - agreement and government. The fact that in surface syntactic structure there are no means other than the labelled syntactic relations to describe interword connections will be essential below.

The surface syntactic structure of (14) is:



(For the surface syntactic relations see Mel'čuk 1974a: 221ff.)

Now the notion 'syntactic zero' can be made explicit. A zero wordform (or zero lex) which has as its signifiant an empty string of phonemes is the syntactic analog of a zero morph. It is the symbol on the nth level of the syntactic representation of a sentence such that it always disappears on movement upward, toward the text; it corresponds to an empty string on level n+1. Conversely, on movement downward, toward the meaning, the zero wordform corresponds to an identifiable complex of symbols on level n-1.

A zero lexeme, a one-element set of lexes which contains only a zero lex, is the syntactic analog of a zero morpheme. The symbol for a zero lexeme, together with the symbols for other lexemes, label the nodes of the syntactic tree on the n' level of syntactic representation.

As was true for zero morphs, a zero lex is deficient only with respect to its signifiant: normally it has a fullfledged signifié and a fullfledged syntactics. It has specific, identifiable meaning and specific, identifiable combinatory possibilities and effects on other words.

A zero lex either belongs to a lexeme that has other, nonzero lexes as well; or it is the sole allolex of a zero lexeme. An example of a zero lex is the zero present tense form of the Russian lexeme BYT' 'be' (copula and existential verb); cf. Evreinov 1973. An example of a zero lexeme is the indefinite personal agent of Russian, discussed

below.

As zero morphemes should be contained in the list of morphemes of a language, so zero lexemes should be contained in the dictionary as separate entries. Zero lexes should also be mentioned in the dictionary entries of their lexemes, just as zero morphs of nonzero morphemes are indicated by the morphological rules. While essentially lexical units, zero lexes and lexemes might also be loosely called syntactic zeroes, in view of the fact that they are introduced and motivated at the syntactic level of linguistic-1 representation. This explains both the title of this paper and my use of the two expressions.

2. This section applies the notions of zero lex and zero lexeme to Russian examples. In (17) and (18)

- (17) Bocman bol'šoj šutnik
 boatswain big joker
 'the boatswain is a great joker'
 (18) Fedor v sosednej komnate
 F. in neighboring room
 'Fedor is in the next room'

the zero lexes Ø-be-1 and Ø-be-2 are normally postulated. These lexes belong to the lexemes BYT' 1 (copula) and BYT' 2 'be located'.

Next consider (19) vs. (20):

- (19) Ulicu zasypali peskom
 street strewed with sand
 acc 3pl past instr
 'the street was strewed with sand'
 (20) Ulicu zasypalo peskom
 strewed
 neut past
 'the street got strewed with sand'

There is a clear difference in meaning. For (19) it is unquestionably people who strewed the street with sand (although just who these people were is not specified), while (20) implies that it was some elemental force (wind, sandstorm, etc.). Thus the meanings of (19) and (20) may be conventionally represented in level n-1 as:

- (19') <people> strewed the street with sand
 (20') <elements> strewed the street with sand

where <people> and <elements> designate the understood agents. These are complex, highly specific meanings distinct from those of the ordinary words people and elements (R. ljudi, stixii). [12, 17]

What expresses the meanings '<people>' and '<elements>' in (19) and (20)? One possibility is that (19) and (20) contain zero nominal wordforms, respectively 0-people and 0-elements, i.e., that these symbols appear on the nth level of sentence representation. It would be natural, then, to label the corresponding nodes in the syntactic structures of these sentences (on level n') with those zero lexemes.

And indeed, not positing zero lexes in (19) and (20), and thus not positing zero lexemes in the syntactic structures of these sentences, leads to difficulties of two types.

Semantic difficulties. If (19) and (20) do not contain zero lexes, and corresponding zero lexemes do not appear in their surface syntactic structures, then evidently the sole source of the meanings '<people>' and '<elements>' would be the verb. But then we would have to admit that almost every Russian verb, taken by itself, is ambiguous in the 3rd person plural and the 3rd person singular neuter. Pisut (3pl) 'they write' would mean either (21) or (22):

- (21) PISAT'-imperf, pres, nonrefl, 3pl
 (22) '<ljudi> pišut', '<people> write'

Similarly, taščilo (neut past) 'it dragged' would be ambiguous between (23) and (24):

- (23) TAŠČIT'-imperf, past, nonrefl, neut
 (24) '<stixii> taščili', '<elements> dragged'

as in (25):

- (25) Ego taščilo po kamnjam, udarjalo o bereg
 him dragged along rocks smashed on shore
 acc neut past neut past
 'he got dragged along the rocks, smashed against
 the shore'

On this analysis supplementary homophonous forms with the meanings '<people>' and '<elements>', and possibly others mentioned below, would be posited for all Russian verbal paradigms. It is bad enough that we would then have to augment every verb paradigm by eight forms: six in the indicative ('<people>' and '<elements>' in the three tenses) plus two in the subjunctive. Moreover, forms like pišut and taščilo are not actually perceived as ambiguous by native speakers; and in fact Russian verbal forms are almost never ambiguous (the single exception is the ambiguity of indicative and imperative second plural forms of a few

second conjugation verbs, unrelated to the present question).

A more serious problem with this analysis is that we would then have to ascribe to grammatical forms (or endings) very complex and specific meanings of the type carried by no other Russian grammatical ending. It is more plausible to postulate two strongly marked and unique lexemes than to posit two homophonous, strongly marked, and unique grammatical endings.

Alternatively, we could say that these indefinite personal and impersonal meanings are present only in certain constructions, namely those where the grammatical subject is materially absent and not recoverable from context. That is, we would claim that the meanings '<people>' and '<elements>' are expressed by the construction as a whole, i.e., by the verb in the given form plus the absence of an overt grammatical subject. But this tantamount to attributing these meanings to the absence of the grammatical subject, which in turn implies that in a formal description of the construction a symbol should be present to point to the absence of grammatical subject. And such a symbol is, in essence, exactly what I am calling a zero lex (functioning as grammatical subject).

Syntactic difficulties. If (19) and (20) lack zero lexes and their surface syntax lacks the corresponding zero lexemes, we have no natural way of accounting for the number, person, and (in the past tense) gender of the finite main verb in these sentences. As has been noted, the morphology of agreement is not expressed in surface syntax. [13] Therefore, in the surface syntax of (19) and (20) the lexeme ZASYPAT 'strew' cannot have indices of person, number, and gender, since these are determined by the grammatical subject and the grammatical subject (on this analysis) is not present.

Rather than resort to ad hoc means, a zero grammatical subject as the source of verb agreement should be posited here, and agreement provided for by the usual mechanism. In (19) the zero grammatical subject is the lexeme \emptyset -people, which triggers 3rd person plural agreement (it is a plurale tantum). In (20) the lexeme \emptyset -elements triggers 3rd person singular neuter agreement (it is a neuter singulare tantum).

Thus the two zero lexemes allow us to easily avoid both of the difficulties. This, in turn, leads to the conclusion that \emptyset -people and \emptyset -elements should appear in a dictionary of Russian, provided with semantic definitions and detailed descriptions of their semantic behavior. [14] The latter point is important, since \emptyset -people and \emptyset -elements have highly specific significés and syntactics, as illustrated below.

First, the Russian '<people>' is more than simply the negation of the signifié '<elements>', and vice versa: the two are not simply logical complements. Neither (26) nor (27)

- (26) Ego vsego iscarapali Ø-people
 him all scratched up
 acc acc 3pl
 'he was all scratched up (by someone)'
 (27) Ego vsego iscarapalo Ø-elements
 neut sing

can refer to animals. For example, neither can be used to describe a situation in which someone is scratched by cats. [15]

Second, Ø-people and Ø-elements have no synonyms in Russian. This means that the corresponding signifiés cannot be adequately and naturally expressed in Russian by any other single means. Curiously, neither Ø-people nor Ø-elements coincides in meaning with French on and il or German man and es. Expressions with on and man are not always translatable into Russian by expressions with Ø-people (see Clas 1970 for English and German translation equivalents to on). The same is true of il and es with respect to Ø-elements.

Third, zero lexemes have limited distribution (possibly due to the nature of their signifiés; much further research remains to be done here). Thus, Ø-people cannot be the subject of a passive: active (28) is correct while passive (29) is not. [16]

- (28) Tam rasstreljali geroev-partizan
 there shot (3pl) heroes-guerrillas (acc)
 'guerrilla heroes were shot there'
 (gram. subj. = Ø-people)
 (29) * Tam byli rasstreljany
 were (3pl) shot 3pl past part.
 palačami
 by executioners
 'there they were shot by executioners'
 (gram. subj. = Ø-people)

Ø-people does not combine with certain verbs, such as naxodit'sja 'be situated', snit'sja 'appear in dreams':

- (30) * Mne vse vremja snjatsja
 me all time appear in dreams (3pl)
 'they (unspecified people) appear all
 the time in my dreams'

Nonetheless, it combines freely with reflexive and reciprocal verbs:

- (31) Tam umyvajutsja
 there wash up (3pl)
 '(people) are washing up there'
 (32) Tam celujutsja
 kiss (3pl)
 '(people) are kissing there'

Presumably this behavior depends on the signifie, which must contain a component 'action' or 'actor' excluding the use of \emptyset -people with passive, 'actionless' predicates.

Fourth, \emptyset -people and \emptyset -elements are marked for case as well as person, number, and gender. However, they can have only the nominative case, since they appear only in the role of grammatical subject, and in Russian the grammatical subject is normally nominative. (Nouns having only one case are attested elsewhere in Russian. Cf. ščec 'some endearing cabbage soup', drovec 'some endearing firewood', drožžec 'some endearing yeast', which have only the partitive.) In much the same way, French on and German man can only be grammatical subjects.

(33) is synonymous with neither (26) nor (27):

- (33) On byl ves' iscarapan
 he was all scratched up
 masc

Nothing in (33) alludes to who or what scratched him. (33) is, however, the syntactic converse of both (26) and (27). The loss of information about the subject in (33) is obligatory because under passivization the demoted subject must appear in the instrumental - which is impossible with a zero lexeme, since it has no instrumental. [17]

Fifth, both \emptyset -people and \emptyset -elements control the gerund just as any ordinary grammatical subject does. It is one of the strictest laws of Russian syntax that a gerund may be used only if it semantic subject (which cannot be overt) coincides semantically and referentially, with the grammatical subject of the governing verb.

- (34) Uvidja nas, on vyšel
 seeing us he went out
 gerund
 'seeing us, he went out'

(35) is a famous jocular example of a typically non-Russian, ungrammatical construction with a 'dangling' gerund:

- (35) * Pod"ezžaja k stancii, u menja
 riding up to station by me
 sletela šlajapa
 fell off hat
 'Riding up to the station, my hat fell off'

Yet with zero subjects gerunds are perfectly acceptable:

- (36) Sjuda kazdyj den' privozjat kirpič,
 here every day bring (3pl) brick
 razgružaja ego u dorogi
 unloading it by road
 gerund
 'Every day they bring bricks here, unloading them by the road'

It is not stated explicitly who brings bricks: it is Ø-people. But those who bring them are the same as those who unload them.

The same is true of (37):

- (37) Liš' v avguste 1539 goda, special'no
 only in August 1539 year, specially
 izmeniv dlja ètogo pravila,
 having changed for this rules
 gerund
 Kardano prinjali v kollegiju vračej
 Cardano accepted into collegium of doctors
 Milana
 of Milan
 'Only in August 1539, having specially changed the rules for it, (they) admitted Cardano to the collegium of doctors of Milan'

Likewise, with Ø-elements:

- (38) Iz èlektrorevolvera xlopnulo, osvetiv
 from electric revolver it cracked lighting
 neut sing
 vse vokrug zelenym svetom (Bulgakov)
 all around with green light
 'From the electric revolver (it) cracked, throwing green light on everything around'

Again, we do not know what produced the cracking noise from the revolver, but it was the same mysterious 'it'

that lighted the environment in green.

Sixth, there is a suggestive parallel between word order adjustments caused by a zero lex, and morphophonemic adjustments caused by a zero morph. The zero morph in Russian cannot carry word stress, for obvious reasons. Thus paradigms with fixed ending stress show automatic retraction of stress onto the stem-final syllable when the ending is zero:

- (39) durák-Ø 'fool' (durak-Ø would be phonetically impossible)
 durak-á
 durak-ú
 durak-á
 durak-é
 durak-óm

In much the same fashion, the zero lex cannot fill a word order position, and another word will automatically be moved to the slot of the zero lex. Russian has a general rule that if the subject is overt and precedes the verb, then the direct object follows the verb:

- (40) Neožidannyj tolčok sbil ego s nog
 unexpected push knocked him from feet
 S V DO
 'an unexpected push knocked him off his feet'

But where no overt subject precedes the verb - where the subject is nonovert or where the subject follows the verb - the direct object is preverbal. (Other orders are possible but marked.) Thus (41), with zero subject:

- (41) Neožidannym tolčkom ego sbilo s nog
 unexpected push him knocked from feet
 instr DO V
 id. (= 40)

Thus the zero subject affects the normal position of the direct object, just as the zero ending affected the normal position of stress in (39).

A syntactic zero lexeme meaning 'something indefinite' has been proposed for Polish and Russian by Wierzbicka 1966, on the basis of the so-called impersonal sentences of both languages. Wierzbicka's zero must serve as grammatical subject in sentences of the type Svetaet '(it) dawns', Morozit '(it) freezes'; in sentences like U menja stučit v viskax 'my temples are pounding', lit., 'by me (it) pounds in temples', Skrebet v glotke '(it) scratches in (the) throat'; and, finally, in sentences like Polja pobilo gradom '(it) crushed the fields with hail', Ego

ranilo oskolkom lit. '(it) wounded him with a shell splinter'.

In addition to Ø-people and Ø-elements, there further appears to be a zero lexeme Ø-any, whose sole (zero) lex has the signifie 'anyone', often in the sense of 'everyone'. Ø-any would be a singular pronoun capable of appearing as the direct object (with verbs governing the accusative case):

- (42) Podobnye poručenija očen' obremenjajut
 such errands very burden
 nom pl
 'such errands are very burdensome',
 lit. '...burden (everyone) very (much)'
- (43) V internate zastavljajut spat'
 in boarding force to sleep
 school 3pl
 posle obeda
 after dinner
 'In boarding school they make (everyone)
 take a nap after dinner'
- (44) Takoe otnošenje očen' raduet
 such attitude very makes happy
 nom sg neut 3sg
 'Such an attitude makes (everyone) happy' [18]

as an oblique object, with verbs governing the dative case:

- (45) Izvestno, čto Zemlja vraščaetsja
 it is known that earth revolves
 vokrug Solnca
 around sun
 'It is known (to everyone) that the earth
 revolves around the sun'
- (46) Nel'zja tak govorit'
 mustn't so talk (inf)
 'No one should talk that way',
 lit. '(one) mustn't...'
- (47) Kurit' vospreščaetsja
 smoke (inf) is forbidden (3sg)
 'smoking is not allowed',
 lit. 'it is forbidden (to all) to smoke'

(Izvestno, nel'zja, and vospreščaetsja govern the dative: mne izvestno, lit. 'to me (it) is known, mne nel'zja 'I shouldn't', lit. 'to me (it's) impossible', mne vospreščaetsja lit. 'to me (it) is forbidden', all with dative mne 'to me'.) Ø-any may also appear as an adnominal, e.g., possessive, modifier. In (48) it is reflexive under

identity to the zero object of tjanet 'burdens':

(48) Svoja noša ne tjanet
 (one's)own burden not burdens
 nom sg 3sg

'Your own load isn't a burden',

lit. '(anyone 's) own load does not burden (anyone)'

Ø-any cannot be the grammatical subject or the object of a preposition. In this respect it resembles the reflexive pronoun sebja 'self', which also has only oblique cases and therefore cannot be the grammatical subject.

3. So far I have postulated three zero lexemes for Russian: the pronouns Ø-people (nominative only), Ø-elements (nominative only), and Ø-any (genitive, dative, accusative). In addition, the verb byt' 'be' presents zero wordforms: the present tense zero lexes Ø-be 1, Ø-be 2, ... This list of zero words and wordforms is undoubtedly opened, and more zero lexemes may prove necessary. Two likely candidates come to mind: Ø-ego in impersonals such as mne xolodno (bol'no, smešno...) 'I'm cold (in pain, amused) = 'my ego experiences cold (pain, humor)'; and Ø-surroundings in impersonals such as zdes' xolodno (grjazno, nakureno, ...) 'here it's cold (dirty, smoky) = 'here the surroundings are cold (dirty, smoky)'. Nothing prevents the discovery of additional zeroes. A number of syntactic zeroes have already been proposed by other investigators; to gain a clearer picture of the relevant constructions I will first provide the following classification of linguistic-2 terminology pertaining to zeroes.

The different uses of the word zero which have appeared in linguistic-2 literature are largely based on Bally 1922 and Jakobson 1938, and can be divided into two classes:

(a) 'zero' as applied to linguistic-1 items: zero phoneme, zero sound, zero affix, zero ending, zero sign, ..., zero article, zero verb, ..., zero predicate, zero grammatical subject, zero grammatical object, ...

(b) 'zero' as applied to entities other than linguistic-1 items: zero paradigm, zero contrast, zero word order, zero stylistic characteristics, zero predicative link, zero valence, ...

As a rule, the uses of type (b) constitute metaphors. They lack a precise common meaning and are replaceable in any particular instance by a different expression. For example, zero paradigm = unmarked paradigm; zero word order = neutral word order; zero valence = absence of valence; etc. Such terminology should be avoided if the meaning of the word zero is not to become completely obscured (cf. Haas 1957: 43, fn. 1).

We must also introduce some order into the uses of type (a) above by assigning a standard and precise meaning to the word zero. In accordance with sections 1 and 2 above I suggest applying the term zero only to the following disjunction: either to linguistic-1 signs (e.g., morphs, lexes) or to sets of synonymous signs distributed according to simple rules (e.g., morphemes, lexemes). Adopting this proposal entails two consequences.

First, terms such as 'zero sound', 'zero phoneme', 'alternation of phoneme /x/ with zero', 'zero signifie', 'zero meaning', etc. are unsatisfactory since sounds, phonemes, and signifiees are neither signs nor sets of signs. In these instances the word zero designates simply 'absence', and thus has a meaning completely different from its meaning in e.g. zero affix. Note that in the morphological representation of a text zero morphs are shown obligatorily (a zero morph is not equal to the absence of a morph), while zero phonemes or phones are never written in phonological or phonetic transcription.

Second, terms such as 'zero predicate', 'zero grammatical subject', 'zero grammatical object', 'zero verb', 'zero noun', 'zero article', 'zero syntactic element', 'zero variant of a word' can be used only to designate zero lexes or lexemes. Thus a zero predicate is a predicate expressed by a zero lex; a zero verb is a zero verbal lexeme; a zero variant of a word is a zero lex of the word, etc.

At the end of section 1 I gave a criterion for evaluating the usefulness of zero signs and corresponding zero '-emes'. That criterion is the investigator's readiness to include the zero sign or zero '-eme' in the same list that contains all similar nonzero signs or '-emes', and to provide the sign or and combinatorial possibilities. Just as a zero affix appears in the inventory of affixes in the language, a zero lexeme, together with its dictionary entry, should appear in the dictionary.

Although this criterion is far from formal, it can obviously be formalized. To do this we need only establish the conditions for defining the willingness or lack thereof of the investigator to include such items in inventories as part of the description of the language. These conditions are probably nothing more than maximal compactness and standardness of inventories; as a rule, linguists try not to enlarge inventories unless it is absolutely necessary. They also seek to avoid duplicating items in such inventories and to avoid grouping unlike items together. The more precise formulation of these conditions is a separate task that I will not deal with here. At present it suffices to strictly correlate the abstract, less intuitively obvious question of the existence of a

zero item to the concrete and much more obvious question of the inclusion of a zero item in the appropriate inventory.

Now we have the apparatus necessary to analyze utterances in the search for syntactic zeroes. I have shown that linguists, when speaking about 'zero X' on the syntactic level, have actually been speaking either of a zero lex or a zero lexeme; and the motivation for postulating a zero lexeme can be verified by the linguist's willingness to include it in the dictionary.

Now let us ask to what extent linguists would be willing to include in dictionaries the 'zero verbs of motion' proposed by Galkina-Fedoruk 1962. She proposes 'zero verbs of motion' as predicates in such sentences as (49). (Here and below, the English words without overt Russian correspondents are capitalized.)

- (49) Tat'jana v les, medved' za nej
 T. into forest bear after her
 'Tat'jana RAN into the forest, the bear
 FOLLOWED her.' (Pushkin)

However, Širjaev (1967, 1973) has demonstrated that carrying this proposal to its logical conclusion would require postulating 'zero verbs of assault' for sentences that are fully analogous to (49) in structure:

- (50) A my ee po tolstym mjasam
 and we her across fat behind
 acc
 'She's GONNA GET IT on her fat behind',
 more lit. 'and we're GONNA GIVE IT TO
 HER on her fat behind'

as well as 'zero verbs of communication':

- (51) Pro svoi deliški on mne ni slova
 about his affairs he to me not word
 'he didn't SAY a word to me about his affairs'

and 'zero verbs of playing':

- (52) My s nej uže vtoruju partiju
 we with her already second game (acc)
 'she and I are already PLAYING the second game'

Evidently there are few, if any, verbal meanings that could not be expressed by such a 'zero verb'. Even diferencirovat 'differentiate' (in the mathematical sense), for example, is replaceable by a 'zero verb':

- (53) A my èto sejčas po t
 and we this now by t
 'Now we will DIFFERENTIATE this by t.'

It follows that if we understand the expression 'zero verb' (of motion, speech, assault, etc.) as a zero lex or zero lexeme we will have the following alternative. Either we include a zero lex (which does not distinguish person, number, or tense) in practically every verbal lexeme; or we introduce a great number of zero word synonyms for almost every verb lexeme in the Russian lexicon. However, those who speak of 'zero verbs' would hardly agree to this. Thus it is clear that in (49-53) there are no 'zero verbs' which could be analogous to the zero nouns \emptyset -people, \emptyset -elements, and \emptyset -any. The word zero in the phrase zero verb is used by Galkina-Fedoruk and Širjaev in a completely different sense than it carries in such phrases as zero affix, zero lexeme.

For (49-53) we would do better to speak of ellipsis (as proposed by Popova 1963: 55-63). The surface syntactic structures of these sentences need to include a top node (as in (16)) labelled by the symbolic name of a specific lexeme. This lexeme carries an essential part of the sentence's meaning; and it determines the form (preposition or case) of the governed noun. Note that case and prepositional government is an exclusively lexical matter: it is conditioned by a lexeme rather than simply by its meaning. Compare the different government patterns of synonymous verbs in (54-55).

- (54) A on nam matematiku
 and he to us (dat) mathematics (acc)
 'he TEACHES us math'
- (55) A on nas matematike
 us (acc) mathematics (dat)
 'he TEACHES us math'

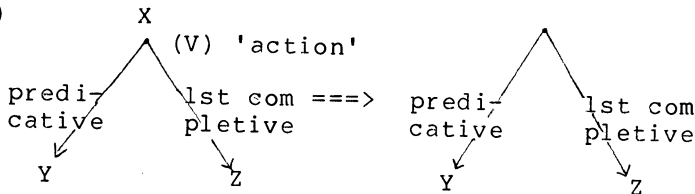
Clearly the surface syntactic structure of (54) must contain prepodavat' 'teach (+Dat +Acc)', while (55) contains učit' or obučat' 'teach (both +Acc +Dat)'. The lexeme in the top node of the surface syntactic structure cannot be zero, as has been established above. And if it is not zero, its government pattern identifies it as a specific lexeme or group of lexemes.

However, on level n-1, i.e., in actual sentences such as (54-55), the top lexeme is not overtly represented. We cannot consider it to be represented by its zero lex: otherwise, as established above, we would have to introduce zero lexes as members of virtually all Russian verbal lex-

emes.

We must therefore accept that in the transition from surface syntactic structure to deep morphological representation the top lexeme is deleted from (54-55). This operation can be described by the following rule.

(56)



Informally stated, in surface syntactic structure the top-most verb, if it denotes an action rather than a state and governs a first complement, may be eliminated; the resulting sentence exhibits a colloquial and expressive character. This rule can apply only to previously morphologized syntactic structures, i.e., the predicate lexeme may be eliminated from surface structure only after the surface markings of the grammatical objects are specified.

Rule (56) is a rule of ellipsis, one of many ellipses possible in Russian. Some details have been omitted from this exposition: the rule is probably not applicable to all verbs designating actions, nor is it always applicable when only a first object is present. For present purposes, however, it suffices simply to characterize ellipsis as a particular type of linguistic rule.

A somewhat different type of ellipsis is observed in imperatives of many languages. In (57-8) the subject is not a zero lex or lexeme, but a deleted second person pronoun.

(57) Stand up!

(58) Behave yourselves!

Sometimes, e.g., under emphasis, the subject can be retained:

(59) You stand up, and you remain seated.

(60) Don't you talk back to your mother!

A very different type of ellipsis appears in (61-3):

(61) Upon arriving, you should go to the passport office.
(= upon your arrival)

(62) I met a friend. (= a friend of mine)

(63) She wants to see the film. (= she wants that she should see the film)

In these sentences there are no identifiable zero lexes, distinct from all other English lexes, which convey specific meanings not attributable to other lexes. Nor can we speak of deletion of surface material that has left syntactic traces in the form of agreement, nonsaturated valences, etc. Rather these involve (obligatory and optional) nonappearance in surface syntax of specific deep syntactic actants.

These examples are far from exhausting the logically possible types of ellipsis; but they suffice to clearly show the difference between syntactic, or lexical, zeroes and ellipses.

Zero and ellipsis, then, are clearly distinguishable, and they contrast in language. Zero is a sign or a set of signs - either a particular sign that has an empty string as its signifiant, or a one element set which contains such a sign. Ellipsis is a rule - one which eliminates certain signs in certain contexts (where they are essentially redundant). Generally, zero conveys meaning, i.e., bears information of some kind. Ellipsis normally does not change meaning but is required by grammatical or stylistic considerations.

Both zero and ellipsis belong to langue rather than parole. Evidence that ellipsis belongs to langue, not parole, is the fact that different languages have different rules of ellipsis. For example, ellipsis of the grammatical subject ja 'I' is stylistically obligatory for Russian performative verbs: prošu vas 'please', lit. 'am asking you'; pozdravljaju vas 'congratulations', lit. 'am congratulating you'. This is not possible in English. Therefore it is incorrect to contrast ellipsis and zero as respectively a phenomenon of parole and a unit of langue (see Bally 1922): both belong to langue. (For zero vs. ellipsis see also Skovorodnikov 1973: 118-9).

While zeroes go into the dictionary, ellipsis is instead included in the grammar among the other syntactic rules (a similar argument for including rules of ellipsis in the grammar is found in Shopen 1972). Syntactic zero and ellipsis as described here do not cover the entire range of phenomena traditionally connected with syntactic incompleteness.

4. I will now answer the questions of Xolodovic mentioned at the beginning of this paper, i.e., I will explain how my analysis can be applied to the problem of grammatical voice in the world's languages.

Question 1. From the viewpoint of voice, (1) contains nothing special:

- (1) Ivana prigasili k trem časam
 I.(acc) invited(3pl) for three o'clock
 'Ivan was invited for three o'clock'

(1) is an ordinary active construction with Ø-people as its grammatical subject. Syntactically, (1) is in no way distinguished from Russian constructions with nonzero grammatical subjects.

Sentences such as (2) are not as straightforward, however.

- (2) Ivana oprokinulo
 I.(acc) knocked over(neut sg)
 'Ivan got knocked over'

If, as proposed here, (2) contains the zero grammatical subject Ø-elements, then we can assign such sentences to the active voice. This is the solution I favor, although the following difficulty must be pointed out: if (2) is active it will be necessary to consider the following (a) and (b) pairs nonsynonymous:

- (64a) Polja pobilo gradom
 fields(acc) crushed(neut sg) by hail
 'the crops were destroyed by hail'
- b) Grad pobil polja
 hail(masc nom) crushed(masc sg) fields(acc)
 'hail destroyed the fields'
- (65a) Glaza rezalo na svetu
 eyes(acc) irritated(neut sg) in light
 '(my) eyes were bothered by the light'
- b) Svet rezal glaza
 light(masc nom) irritated(masc sg) eyes(acc)
 'the light bothered (my) eyes'
- (66a) Sil'nyj udarom ego
 by strong(instr) blow(instr) him(acc)
 sbilo s nog
 knocked off(neut sg) from feet
 'he got knocked off his feet by a strong
 blow'
- b) Sil'nyj udar sbil
 strong blow(masc nom) knocked(masc sg)
 ego s nog
 him(acc) from feet
 'a strong blow knocked him from his feet'

The (a) sentences must be understood as '<elements>

crushed the fields with hail', etc., and the (b) sentences as 'hail crushed the fields', etc. In other words, the relationship between the (a) and (b) sentences in (64-66) is the same as that between (67a) and (67b):

- (67a) Karandaš provel tonkuju
pencil(masc nom) drew(masc sg) thin(acc)
liniju
line(acc)
'the pencil drew a fine line'
- b) Ivan provel karandašom tonkuju
I.(nom) pencil(instr)
liniju
'Ivan drew a fine line with the pencil'

where (67b) has an additional overt actant corresponding to '<elements>' of (64-66b).

I am prepared to accept the interpretation of the pairs as nonsynonymous. (64a-b) would appear to be a counterexample, since they are perceived by native speakers as semantically identical. However, since hail is an 'element' itself, both of these (nonsynonymous) sentences have one and the same real world referent and therefore seem to have identical signifiés. That (66a-b) are not synonymous is more readily apparent: (66a) indicates that what struck him was something unclear or incomprehensible, while there is no such meaning in (66b). This difference underlies the unacceptability of (68) vs. the acceptability of (69):

- (68) Soldaty brosilis' na Ivana, i srazu že
 soldiers rushed at I. and immediately
 *sil'nym udarom ego sbilo
 strong blow(instr) him(acc) knocked(neut sg)
 s nog
 from feet
 'the soldiers rushed at Ivan and immediately he
 got knocked off his feet by a strong blow'
 (69) Soldaty brosilis' na Ivana, i srazu že
 sil'nyj udar sbil ego s nog
 (masc nom) (masc sg)
 'the soldiers rushed at Ivan and immediately
 a strong blow knocked him off his feet'

Likewise, the nonsynonymity of the second clauses is evident in (70-71):

- (70) Ivan otkryl kran, i *vodoj
 I. (nom) opened faucet and by water (instr)
 srazu že zalilo plastinu
 immediately flooded photographic plate
 (neut sg) (acc)
 'Ivan turned on the faucet and the plate was
 immediately flooded with water'
- (71) Ivan otkryl kran, i voda srazu že
 (fem nom)
 zalila plastinu
 (fem sg)
 'Ivan turned on the faucet and water immediately
 flooded the plate'

The alternative to this solution is to consider the (a) and (b) pairs synonymous. This would entail positing for Russian, and assigning to (2) and (64-66a) a special voice category, 'impersonal' or 'subjective impersonal'. Then the forms oprokinulo in (2), pobilo in (64a), and rezalo in (65a) would cease to be personal forms requiring agreement. The dummy subject would no longer be needed to provide for their morphological shape, since they would simply be impersonal voice forms that lack personal conjugation (much like the Estonian impersonal voice of (77-78) below). There are at least two unpleasant consequences to this alternative solution. First, all third person singular neuter verb forms in Russian would become ambiguous, between '3sg, neut, indicative' and 'impersonal'. Second, there would no way to explain the ungrammaticality of (68) and (70).

Question 2. (1-4) are not subjectless sentences: each contains a grammatical subject in the form of a zero lexeme which figures in their surface syntactic structures but does not materialize in the transition to the actual sentence. Moreover, for Russian and all other languages where the main verb obligatorily agrees with the grammatical subject (Spanish, English, etc.) there can be no subjectless finite sentences at all. Elliptical sentences are of course possible. If, however, a sentence contains a finite verb, the verb must agree with something, and this can only be the grammatical subject, including one expressed by a zero lexeme. The following Russian sentences are bipartite, i.e., they consist of (zero) subject + VP:

- (72) Cypljat po oseni sčitajut
 chicks(acc) in fall count(3pl)
 'chicks are counted in the fall'
 (a proverb, roughly 'don't pass
 judgment prematurely')
- (73) Morozit
 freeze(3sg)
 'it is freezing'

much as Russian lexes such as ruk 'of hands' (gen pl), nos 'nose' (nom sg) are bimorphemic. In light of this, (73) is not subjectless; it is, however, impersonal, since its grammatical subject can only be \emptyset -elements. [19]

Similarly, (3) is not subjectless, since it is also bipartite:

- (Ukrainian)
 (3) Bulo organizovano ekspedyciju
 was(neut sg) organized(neut sg) expedition(acc)
 'an expedition was organized'

It has a zero grammatical subject (neuter zero pronoun) plus a predicate (neuter singular, in agreement with the subject) with its direct object. From the viewpoint of voice, (3) is active. More specifically, it has subjective impersonal voice (see Mel'čuk & Xolodovič 1970: 118), where the meaning of the human agent is imparted by the verbal form in -to/-no (cf. zavezeno '(they) have brought'). The Spanish example (4) is also a subjective impersonal, rather than a passive, construction.

Note that constructions such as (3) and (4) contain an unusual kind of zero grammatical subject. It is a semantically empty zero lexeme, i.e., a zero lexeme with an empty semantic graph for a signifié. Empty lexemes are attested in many languages: strongly governed prepositions and conjunctions, various filler words (e.g. the Persian izafet), and 'dummy' ('grammatical', 'expletive') subject pronouns. Examples of the latter are Eng. it, Fr. il, Ger. es as in:

- (74) It is evident that the parser shouldn't...
 (75) Il s'agit ici de trois difficultés suivantes
 lit. 'it is dealt here with three following
 difficulties'
 (76) Es ist hier viel Ski gefahren
 'people ski here alot', lit.
 'it is skied here very much'

It in (74) has the signifiant /it/ and a very complex syn-

tactics, but its signifié is an empty set of semes:

it = </it/; 'Λ'; Σ = anticipatory pronoun in such constructions as...>

The same holds of Fr. il and Ger. es in (75-76). The Ukrainian dummy grammatical subject in (3) is likewise empty; but unlike it, il, es it is also materially zero. The only component spared this mutilated pronoun is its syntactics: it is singular neuter as shown by the verb agreement, it can be used only in a strictly limited type of construction, etc.

The mere presence of syntactics as manifested in agreement is sufficient basis for postulating a zero lex and corresponding lexeme. Such signs, simultaneously empty and zero, represent a degenerate case and are rarely encountered. [20]

It should be emphasized that no zero lex subject may be postulated where the verbal form does not exhibit agreement. Thus in Estonian forms in -takse (subjective impersonal voice) are not inflected for person or number at all:

- (77) Kuulatakse muusikat
is listened to music(acc)
'music is listened to'
- (78) Haalikut [o] haaldatakse nii
sound(acc) is pronounced so
'the sound [o] is pronounced like this'

No zero lex can be posited here.

Question 3. A zero agentive complement with indefinite personal meaning is not postulated for sentences such as (5)

- (5) Ivan byl priglašen k trem časam
I. was invited for three o'clock
(nom) (masc sg) (masc sg)
'Ivan was invited for three o'clock'

because the meaning '<people>' is not expressed in such constructions (see the discussion of (33)). Likewise, in (33) and (79) there is no reference whatsoever to who or what scratched him or knocked him over - it could be people, the elements or animals.

- (33) On byl ves' iscarapan
he was(masc sg) all scratched up(masc sg)
'he was all scratched up'

- (79) On byl oprokinut
knocked over (masc sg)
'he was knocked over'

(79) is not synonymous to either (80) or (81):

- (80) Ego oprokinuli
him(acc) knocked over(3pl)
'he was knocked over (by people)'

- (81) Ego oprokinulo
(neut sg)
'he got knocked over (by the elements)'

In addition to the lack of semantic basis, there is also no natural syntactic basis to justify a 'zero agent' in such constructions as (5), (33), and (79): the presence of such an agent would never be manifested in surface phenomena such as agreement. It is true that the agent is perfectly recoverable in (5), and that it is identifiable as 'people', i.e., (5) is synonymous with (1):

- (1) Ivana priglasili k trem časam
I.(acc) invited(3pl) for three o'clock
'(they) invited Ivan for three o'clock'

However, this synonymy depends on the particular meaning of the verb priglašat 'invite' and on our knowledge of the real world. Insofar as only people (including people in a broad sense - collectives, institutions, organizations, etc.) can be engaged in the activity of inviting, the meaning 'people' (as agent) in (5) emerges from the verb priglašat. This accounts for the synonymy of (1) and (5). As we have seen, verbs that do not describe specifically human actions lack such synonymy (see (65) and (66) above).

(Translated from Russian by Mark Green, Cornell University)

FOOTNOTES

- * This paper is a revised and enlarged translation of Mel'čuk 1974b. I wish to thank Ju.D. Apresjan, A.Ja. Dikovskij, L.L. Iomdin, L.N. Iordanskaja, E.N. Širjaev, A.A. Xolodovič, and A.K. Žolkovskij, who read the initial drafts of this paper and pointed out a number of errors. A particular debt of gratitude is owed to Susanne Carroll, my first North American reader and

critic; L.N. Iordanskaja, who went through the final version pinpointing inconsistencies and obscure statements; and Johanna Nichols, who edited it. Needless to say, I am alone responsible for surviving drawbacks and blunders.

- 1 Aleksandr Aleksandrovic Xolodovič, one of the leading Russian theoretical linguists and a brilliant Japanologist and Koreanologist (his partial bibliography is found in *Narody Azii i Afriki*, 1966: 3. 215-17; 1976: 6. 229) died of heart failure in 1977. For more than twenty years he had been my linguistic guru and, more important, a friend. May the present paper be a modest contribution to preserving his memory.
- 2 The English adjective linguistic is ambiguous between 'pertaining to language(s)' (Ger. sprachlich) and 'pertaining to linguistics' (Ger. sprachwissenschaftlich). Since in this paper the distinction is crucial, I will use the following: linguistic-1 denotes the first sense, linguistic-2 the second.
- 3, Elementary X = X which cannot be represented in terms of other X's.
- 4 Thus, in my use of the term a morpheme (or a lexeme - see below) is not a sign but a set of signs.
- 5 For insightful remarks about zero suffix morphs in English see Smirnitskij 1959: 20-3. As regards the much discussed zero morpheme of singular in English nouns, I think it is a zero morph but not a zero morpheme, since beside zero (as in house-Ø, bird-Ø, leg-Ø) we also find some nonzero morphs (phenomen-on, alumn-us, formul-a). Although the latter are few in number and found only in Latin borrowings (most in carefully written texts), I feel their presence justifies postulating a nonzero morpheme for English: {SG} = -Ø-sg, -on, -us, -a,....
- 6 Zaliznjak (1967: 19) uses word segment for 'signifiant of a wordform'.
- 7 In section 4 we will see that in principle neither justification - justification from below alone, or justification on the same level alone - is sufficient grounds for postulating a zero sign. In fact, empty zero signs with only the syntactics nonempty are possible. Conditions on postulating zero forms are discussed at length and clearly in Haas 1957.
- 8 See in this connection the wellknown paper by Benveniste 1964.
- 9 Generalizing Zaliznjak's concept (fn. 6), we might say sentence segment.
- 10 A lex is a (grammatical) form of a lexeme; it can be either a wordform or a phrase representing an analytical form of the lexeme in question. some lexes of the English lexeme SEE: see, sees, will see, was seen,....

- 11 na-1 is the preposition na 'onto' governing the accusative case, as opposed to na-2 'on', which governs the prepositional case.
- 12 For English, something may be a more felicitous rendition of the second zero than elements, the literal translation of Russ. stixii. This is also what Wierzbicka suggests (1966: 188,191,193).
- 13 I cannot justify such a surface syntactic structure here. I will only point out that insofar as the linguist's surface syntax has as its goal an explicit and homogeneous description of the syntactic makeup of actual sentences, then purely syntactic devices such as syntactically conditioned prosody, word order, and syntax oriented morphology (agreement, government) should not be depicted in surface syntactic structure, tels quels, as a motley, confusing, and ambiguous set. They should be represented more abstractly and formally by means of surface syntactic relations, designed by the analyst in much the same way transcriptions are designed to represent actual sound. Thus the syntactic devices in and of themselves are not preserved in the surface syntax of the actual sentence. For more on surface syntactic relations, syntactic trees, etc. see Mel'čuk 1979.
- 14 This conclusion fully coincides with the opinion of Panov (1960: 11): 'the grammatical subject can be zero, there being in Russian several homonymous zero grammatical subjects, as in svetaet '(it) dawns', cypljat po oseni sčitajut '(they) count chicks in the fall', etc. Cf. corresponding nonzero subjects in other languages, e.g., Ger. es, man.
- 15 Some complexities, however, remain, for further investigation. Consider such normal sentences as the following:

- (i) Nado že, vse sklevali
it's incredible everything pecked up(3pl)
'I can't believe everything got pecked up',
lit. '...(they) pecked everything up'
- (ii) Bednen'kij, kak tebja pokusali!
poor thing how you(acc) bit(3pl)
'Poor thing, you got all bitten up!'
lit. '(they) bit you all up'

In these sentences the zero grammatical subject would designate birds and insects, respectively. Compare the following unacceptable sentences:

- (iii) *Takoј silos ne edjat
 such sileage(acc) not eat(3pl)
 'such sileage (they) don't eat'
 (where (they) = cows)
- (iv) *U nas pasutsja na bol'som lugu
 at us graze(3pl) in big pasture
 'we have (them) grazing in the big pasture',
 '(they)'re grazing...'

Nonovert subjects can be understood to refer to birds or insects, but not cattle. It is not unlikely that there is ellipsis of the grammatical subject, rather than a zero lexeme, in (i) and (ii).

- 16 Sentence (29) may be correct if read as elliptical, i.e., the result of deleting the grammatical subject oni 'they' by coordinate reduction:

- (i) Tam i byli sxvačeny geroi-partizany.
 there were captured heroes-guerrillas
 (nom pl)
 Tam byli rasstreljany palačami.
 'It was there that the guerrilla heroes
 were captured. There they were shot by
 executioners'

- 17 For interesting data about the semantics and syntax of \emptyset -people see Nakhimovsky 1978, esp. Ch. 3 ('Syntactic zeroes'). His claims may be roughly summarized as the two relevant properties of this zero lexeme.

First, \emptyset -people is semantically much like the Russian indefinite pronouns kto-nibud', kto-to 'someone', etc. For instance, \emptyset -people forces the epistemic reading on what would elsewhere be ambiguous sentences with mož' 'can, may', just as indefinite pronouns do.

- (i) On možet priglasit' Dimu k obedu
 he can/may invite D. for dinner

(i) means either 'he is physically able/allowed to invite Dima to dinner' or 'it is possible (it may happen) that he will invite Dima to dinner'. With \emptyset -people or kto-nibud' in an independent clause only the epistemic reading is possible:

- (ii) Kto-nibud' možet priglasit' Dimu k obedu
 someone
- (iii) Dimu mogut priglasit' k obedu
 can/may(3pl)

Both (ii) and (iii) can be understood only as 'it is possible that someone will invite Dima to dinner', or (for (iii)) 'Dima just may get invited to dinner'. In this respect Russ. \emptyset -people seems to be in sharp contrast with Fr. on, Ger. man.

Second, \emptyset -people may refer not only to many people but also to a single person, including the speaker or the hearer:

- (iv) Utrom že tebe žrat' davali...
 in morning you(dat) to eat gave(3pl)
 Ili net, èto ja tebe včera daval
 or it(s) I yesterday gave
 'You were fed in the morning, weren't you...
 Actually you weren't - it was yesterday that
 I fed you' (man speaking to his cat; quoted
 from Nakhimovsky 1978: 109)

However, reference to the speaker is excluded in contexts like the following:

- (v) Ja obeščaju, čto Dimu priglasjat
 I promise that D. will invite(3pl)
 k obedu
 for dinner
 'I promise that Dima will be invited to dinner'

(v) can only mean that someone other than the speaker with invite Dima. The same is probably true of Fr. on and Ger. man.

A fairly exhaustive review of the various uses of the Russian indefinite personal construction is offered by Gasparov (1971). Many interesting examples and penetrating remarks on the meaning and contrastive use of Fr. on, Ger. man, and Eng. one, anybody, people and the passive construction are found in Clas 1970.

- 18 This allows us to consider the direct objects of the transitive verbs of (42-44) as strictly obligatory. The nonovert status of the direct object conveys specific information and is thus treated as a zero lex of the zero lexeme \emptyset -any. This is not true of verbs such as čitat 'read', verit 'believe', pet 'sing', and others: for these the direct object is syntactically facultative, and its absence conveys no specific information. In the following sentence nothing is said about what he is reading:

- (i) On sidel i čital
 he sat and read
 'he sat and read'

(Cf. also in this connection Lehrer 1970).

- 19 Interestingly enough, all the claims made in this paragraph were stated in completely explicit form as early as 1935 by Jakobson (1935 (1971): 21): 'The Russian norm does not know personal sentences without grammatical subject. The so-called impersonal sentences exhibit a zero subject. Russian has lost all types of unipartite narrative sentences.' Jakobson even offers a diachronic explanation, relating the appearance of the zero lexeme subject in Russian to the disappearance of inflected enclitics in all the northeastern Slavic languages. Cf. also the insightful remark of Hetzron (1969: 141): 'In Italian piov-e 'it rains' the grammatical subject is zero rather than simple absence, since its presence is signalled indirectly by the suffix of the verb [3sg - IAM]'.
- 20 Another example of an empty zero lex is reported by Hetzron (1969: 152-3): his zero object in some Hungarian idiomatic constructions:

- (i) X megjár-ja Y-kel 'X has trouble with Y'
- (ii) X beér-i Y-kel 'X satisfies himself with Y'
- (iii) X megér-i Y-nek 'X is worth doing for Y'

Here all the verbs are in the so-called objective conjugation, shown by the hyphenated 3sg object suffix. The Hungarian objective conjugation ordinarily indicates agreement in number and person with a definite direct object, but in these constructions no direct object is overtly present. The verb agreement leads Hetzron to postulate a zero direct object - a dummy definite noun that is semantically void. It does not contribute any specific meaning to the above expressions, yet its presence is crucial for their idiomatic meaning: megjár means 'pass by', while megjár + \emptyset direct object means 'have trouble'; beér means 'catch up with', while beér + \emptyset direct object means 'satisfy oneself with'; megér means 'cost', while megér + \emptyset direct object means 'be worth doing'.

Constructions with a zero empty subject, similar to (3), are typical of north Russian dialects:

- (i) Molodu ženščinu sxvačeno medvedicej
 young woman taken by she-bear
 (acc) (pass part neut)
 'A young woman has been carried off by a
 she-bear'

- (ii) Babuŝku -to gde poxoronenno?
 Grandma ptc where buried
 (acc) (pass part neut)
 'where is Grandma buried?'
- (iii) Otravleno bylo syna
 sent off was son
 (pass part neut) (neut) (acc)
 '(my) son was sent away'

(quoted from Babby & Brecht 1975: 347). For all these sentences a zero empty subject (functionally equivalent to Eng. expletive *it*) may be posited to account for the neuter gender of the predicate. For more such examples see Kuz'mina & Nemčenko (1971: 27-106). This is not the only possible analysis. Another approach is suggested by Timberlake (1976): in the following sentences the prepositional phrase u + noun is taken as subject:

- (iv) U menja bylo telenka zarezano
 at me was calf slaughtered
 (neut) (acc) (neut)
 'I have slaughtered a calf'
- (v) Vodu u ej naneseno
 water(acc) at her brought(neut)
 'she has brought water'

The agreement rule is as follows: if the grammatical subject is the phrase u + noun, then the main verb is in the neuter gender. But too many things remain unclear to pass definitive judgment.

REFERENCES

- Aschmann, H. and W.L. Wonderly. 1952. Affixes and implicit categories of Totonac verb inflection. *IJAL*, 18.3. 13-45.
- Babby, L.H. and R.D. Brecht. 1975. The syntax of voice in Russian. *Lg.*, 51.2. 342-67.
- Bally, Ch. 1922. Copule zéro et faits connexes. *Bulletin de la Société de Linguistique de Paris*, 23. 1-6.
- Barxudarov, L.S. 1966. *Struktura prostogo predloženiya sovremennogo anglijskogo jazyka*. Moscow.
- Benveniste, E. 1964. Les niveaux de l'analyse linguistique. *Proc. of the 9th Intl. Congress of Linguists*. 266-75. London, etc.: Mouton.

- Clas, A. 1970. Le système du pronom indetermine on: problemes de traduction. *Lebendige Sprachen*, 25.1. 13-6.
- Evreinov, Irina A. 1973. Die Semantik einer Nullform: Versuch einer neuer Definition der Kopula im Russischen. *Linguistics*, 98. 41-57.
- Galkina-Fedoruk, E.M. 1962. O nulevyx formax v sintaksise. *Russkij jazyk v škole*, 2. 6-12.
- Gasparov, B.M. 1971. Neopredelenno-subektnye predloženiya v sovremennom russkom jazyke. *Trudy po russkoj i slavjanskoj filologii*, XIX, Tartu University. 3-58.
- Haas, W. 1957. Zero in linguistic description. In: *Studies in linguistic analysis*. 33-53. Oxford: Basil Blackwell.
- Hetzron, R. 1969. Des compléments obligatoires en hongrois. *Word*, 25.1-2-3. 140-54.
- Jakobson, R.O. (1935) 1971. Les enclitiques slaves. *Selected writings*, II. 16-22. The Hague-Paris: Mouton.
- (1939) 1971. Signe zéro. *Selected writings*, II. 211-9. The Hague-Paris: Mouton.
- Kastovsky, D. 1969. Wortbildung und Nullmorphem. *Linguistische Berichte*, 2. 1-13.
- Kuz'mina, Irina B. and Elena V. Nemčenko. 1971. Sintaksis pričastnyx form v russkix govorax. Moscow: Nauka.
- Lehrer, Adrienne, 1970. Verbs and deletable objects. *Lingua*, 25.3. 227-53.
- Lopatin, V. 1966. Nulevaja affiksacija v sisteme russkogo slovoobrazovanija. *Voprosy jazykoznanija*, 1. 76-87.
- Meier, G.F. 1961. Das Zero-problem in der Linguistik. Berlin: Akademie-Verlag [Schriften sur Phonetik, Sprachwissenschaft und Kommunikationsforschung, 2.]
- Mel'čuk, I.A. 1967. K postroeniju dejstvujuščej modeli jazyka. In: *Problemy jazykoznanija*. 82-9. Moscow: Nauka.
- 1972. Urovni predstavlenija vyskazyvanija i obščee stroenije modeli "Smysl \Leftrightarrow Tekst". Moscow: Institut Russkogo Jazyka Akad. Nauk SSSR.
- 1974a. Opyt teorii lingvističeskix modelej "Smysl \Leftrightarrow Tekst". Moscow: Nauka.
- 1974b. O sintaksičeskom nule. In: *Tipologija passivnyx konstrukcij. Diatezy i Zalogi*. 343-61. Leningrad: Nauka.
- 1976. *Das Wort*. Munchen: W. Fink Verlag.
- 1979. *Studies in dependency syntax*. Ann Arbor (Mich.): Karoma.

- and A.A. Xolodovič. 1970. K teorii grammatičeskogo zaloga. (Opredlenije. Isčislenije). Narody Azii i Afriki, 4. 111-24.
- Nakhimovsky, A. 1978. Voice and the indefinite-personal construction in Russian. Ph.D. thesis. Cornell University. Ithaca, N.Y.
- Panov, M.V. 1960. O častjax reči v rusckom jazyke. Filologičeskije nauki, 4. 3-14.
- Popova, I.A. 1963. Npolnyje predloženiya v sovremennom rusckom jazyke. Trudy Instituta Jazykoznanija Akademii Nauk SSSR, II. 3-136.
- Shopen, T. 1972. Ellipsis as grammatical indeterminacy. Ms.
- Skovorodnikov, A.P. 1973. O kriterii èlliptičnosti v rusckom sintaksise. Voprosy jazykoznanija, 3. 114-23.
- Smirnickij, A.I. 1959. Morfologija anglijskogo jazyka. Moscow: Izdatel'stvo Literatury na Inostr. Jazykax.
- Širjaev, E.N. 1967. Nulevyje glagoly kak členy paradigmaticeskix i sintagmaticeskix otnošenij (na materiale sovremennogo rusckogo jazyka). Avtoref. kand. diss., Moscow.
- , 1973. O nekotoryx pokazateljax nezameščennyx pozicij v vyskazyvanijax razgovornoj reči i nulevyje glagoly-predikaty. In: Russkaja razgovornaja reč'. Moscow: Nauka.
- Timberlake, A. 1976. Subject properties in the North Russian passive. In: Ch. Li and S. Thompson (eds.). Subject and Topic. 545-70. New York, etc.: Academic Press.
- Vardul', I.F. 1969. K voprosu o javlenii èllipsisa. In: Invariantnyje sintaksičeskije značeniya i struktura predloženiya. 56-70. Moscow.
- Wierzbicka, Anna. 1966. Czy istnieją zdania bezpodmiotowe. Język polski, 46.3. 177-96.
- Zaliznjak, A.A. 1967. Russkoe imennoe slovoizmenenie. Moscow: Nauka.
- Žolkovskij, A.K. and I.A. Mel'čuk. 1967. O semantičeskom sinteze. Problemy kibernetiki, 19. 177-238.
- , 1969. K postroeniju dejstvjuščej modeli jazyka "Smysl \Leftrightarrow Tekst". Mašinnyj perevod i prikladnaja lingvistika, 11. 5-35.